DOCUMENT RESUME

ED 320 121 CS 010 116

AUTHOR Valencia, Sheila; And Others

TITLE Assessing Reading and Writing: Building a More Complete Picture for Middle School Assessment. Technical Report No. 500.

INSTITUTION Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.

SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE Jun 90
CONTRACT G0087-C1001-90
NOTE 36p.
PUB TYPE Guides - Non-Classroom Use (055) -- Guides - Classroom Use - Guides (For Teachers) (052)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Evaluation Methods; Intermediate Grades; Junior High Schools; *Literacy; Middle Schools; *Portfolios (Background Materials); *Reading Achievement; Reading Writing Relationship; Self Evaluation (Individuals); *Student Evaluation; *Writing Evaluation
IDENTIFIERS *Middle School Students; Student Centered Assessment

ABSTRACT

The intention of this report is to build a strong case for a portfolio approach to assessment of reading and writing during the middle school years. A portfolio is both a collection of the artifacts of students' reading and writing and a disposition toward assessment. This disposition is characterized by the belief that assessment must be continuous, multidimensional, collaborative, knowledge-based, and authentic. Recommendations for portfolios are based upon the conviction that records gathered while students are engaged in functional and contextualized literacy tasks will ultimately prove more useful to both teachers and students than will any set of numbers derived from tests that have little relevance to the purposes and needs of either students or teachers. The principles and continua put forward should be regarded as a conceptual framework to use in developing and evaluating assessment strategies. The array of assessment opportunities available to teachers is limited only by their creativity, knowledge of the reading process, and grasp of sound assessment criteria. The report describes the portfolio assessment practices in detail and illustrates how they must be applied in the classroom by means of five "scenarios." Appendixes include: A Schema for Scoring Retellings (The Retelling Profile), A Wholistic Scoring Scale, An Analytic Scoring Scale, A Writing Analysis Checklist, and A Self-Evaluation Questionnaire. One figure and 53 references are included. (MG)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED320121

CENTER FOR THE STUDY OF READING

Technical Report No. 500

ASSESSING READING AND WRITING: BUILDING A MORE COMPLETE PICTURE FOR MIDDLE SCHOOL ASSESSMENT

Sheila Valencia
University of Washington

William McGinley
University of Michigan

P. David Pearson
University of Illinois at Urbana-Champaign

June 1990

University of Illinois at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

D. ANDERSON

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

CS010116

The work upon which this publication was based was supported in part by the Office of Educational Research and Improvement under Cooperative Agreement No. G0087-C1001-90 with the Reading Research and Education Center. The publication does not necessarily reflect the views of the agency supporting the research.

BEST COPY AVAILABLE

Abstract

The revolution in literacy assessment in the past half-decade has raised serious concerns about conventional assessments. We go beyond those concerns to develop a framework. Specifically, we advocate a portfolio approach to assessment, then explain and illustrate the approach with classroom examples from the middle school level.

EDITORIAL ADVISORY BOARD
1989-90

James Armstrong
Linda Asmussen
Gerald Arnold
Yahaya Bello
Diane Bottomley
Catherine Burnham
Candace Clark
Michelle Commeyras
John M. Consalvi
Christopher Currie
Irene-Anna Diakidoy
Barbara Hancin
Michael J. Jacobson
Jihn-Chang Jehng
Robert T. Jimenez
Bonnie M. Kerr
Paul W. Kerr
Juan Moran
Keisuke Ohtsuka
Kathy Meyer Reimer
Hua Shu
Anne Stallman
Marty Waggoner
Janelle Weinzierl
Pamela Winsor
Marsha Wise

MANAGING EDITOR
Fran Lehr

MANUSCRIPT PRODUCTION ASSISTANTS
Delores Plowman
Debra Gough

ASSESSING READING AND WRITING: BUILDING A MORE COMPLETE PICTURE FOR MIDDLE SCHOOL ASSESSMENT

This report is a plea for putting the responsibility for and control of literacy assessment back into the hands of those whose lives are most affected by it--students and teachers. It is organized into five parts. We begin with a rhetorical justification for our position on classroom control of assessment. We introduce the attributes that characterize good classroom assessment. Then we present an overall framework for examining various aspects of the reading and writing processes. The heart of the report presents a set of constructs and a language for describing activities that might become part of a portfolio. Finally, we illustrate our perspective in a set of scenarios, examples of processes used by middle school teachers to gather data for student portfolios.

For students, the middle school years represent a period of transition concerning the way in which they must come to view reading and writing. While fluent reading and writing continue to be major goals of their literacy instruction, middle school students must also learn how to use reading and writing as ways of learning and communicating in all subject areas and in daily life. Anderson, Hiebert, Scott, and Wilkinson (1985) aptly describe the role of reading in the challenge of subject matter learning: "As proficiency develops, reading should be thought of not so much as a separate subject in school but as integral to learning literature, social studies, and science" (p. 61).

Several researchers have documented a similar role for writing as an means of fostering learning and understanding in subjects such as literature, social studies, and science (Applebee, 1981, 1984; Gage, 1986; Langer & Applebee, 1987; Martin, Meday, Smith, & D'Arcy, 1975).

If the middle school years represent a unique period of instructional transition during which reading and writing become the primary means by which students can learn or accomplish other goals, then assessment in a middle school literacy curriculum must reflect these functions. Teachers and researchers must begin to explore assessment procedures that allow them to examine students' ability to use reading and writing as tools for learning, insight, understanding, communication, enjoyment, survival, and participation in their world.

Historically, both the general public and professional educators have come to regard assessment as a series of paper and pencil tests imposed on teachers and students from those in positions of authority--those concerned with accountability. All too often, assessment has been something we do for "them," the authorities, rather than something we do for ourselves and our students. However, assessment *is* (or ought to be) what we do to help ourselves and our students evaluate progress. It is through daily interactions that we gather evidence to guide decision making about student learning, instruction, and curriculum. Assessment should be a natural part of the teaching/learning process--not something added on or imposed as an afterthought.

Teachers and administrators need to learn to rely on their own expertise to develop and use a wide variety of assessment tools to draw conclusions about student progress. We have reached a stage in the development of our knowledge about reading and writing processes at which we can no longer rely solely upon formal and indirect performance measures, such as those found on most commercially available tests--the norm-referenced standardized tests or criterion-referenced basal tests so often associated with assessment. In arguing for a return to measures that are grounded more directly in classroom instruction, we do not advocate that educators discontinue using commercially available tests. Instead, we argue for a more realistic perspective on the value of commercially available tests, one that admits what we know to be true about all educational decision making: that it is dangerous to rely on any single score as a measure of learning, no matter how scientifically sound it may appear to be.

As teachers, our experience and intuition (what Shulman [1975] has called the practical wisdom of teaching) tell us that student evaluation requires more than one simple score on a paper and pencil measure. We know that we can learn much about students by keeping track of their learning during classroom interactions. Our assessment repertoire must be expanded to include many more exemplars of what some have come to call "situated" (Collins, Brown, & Newman, 1989) or contextualized assessment strategies (Lucas, 1988a; Lucas, 1988b); we need indicators of learning that are gathered as an integral part of the teaching/learning situation.

To the commonly used decontextualized measures such as formal standardized and criterion-referenced tests, we must add assessments of situated performances--ongoing analyses of students completing a variety of literacy tasks that have been initiated at the prerogative of the school, the teacher, or the student. Performance measures alone, even when they are personally initiated and functionally motivated, are not sufficient. Equally important are the kinds of information that we can gather only by observing students during literacy events or holding conversations with them to determine what they think about their own skills, strategies, and dispositions toward literacy.

Important Attributes of Classroom and Individual Assessment

As we move to more and more contextualized assessments of reading and writing, the pool of assessment opportunities expands exponentially. In every classroom, every day, teachers can, do, and should use many different strategies to help them build a more complete picture of students' literacy. Later in this report we will discuss several assessment options; first, however it is essential to describe the guiding principles, the attributes, of sound contextualized assessment. Because it is possible to present only a few examples from a seemingly limitless pool of assessment strategies, these principles can serve as important criteria for teachers to use to evaluate, design, and characterize their own assessment strategies. That is perhaps the most exciting and gratifying aspect of this entire endeavor--the potential for teachers to derive assessments that reflect their students, classrooms, curricula, and themselves. The principles:

1. **Assessment is continuous**. Learning is a continuous, dynamic process. Not only does learning take place over time, but the learner and task change with every new instructional situation. Although we frequently acknowledge this idea conceptually, we rarely apply it to classroom evaluation. Unfortunately, the framework within which most evaluation is based is similar to the common notion of end-of-unit or end-of-year tests. That is, we have come to regard classroom assessment as a static entity, as a test or set of tests that are administered to students and then graded by a teacher to measure what the student has learned about some topic, skill, or strategy.

There are several risks in failing to adopt a mind set for continuity; we may forget (a) to look closely at learners' growth, (b) that knowledge is cumulative and transferable, and (c) that the basis for future learning comes precisely from growth in reasoning ability and integration of knowledge over time.

We offer a caution, however, for those who are overzealous in their affinity for structure and consistency. These ongoing measures we advocate do not need to be staged--they need not be formal or to result in a grade. In fact, their beauty lies in their simplicity and ubiquity; they allow us to take advantage of what happens daily in most classrooms. The objective is to inform instruction and to provide feedback and support to students. We would be wise to take heed from Diederich (1974) as he discusses his impressions of assessment in writing classes:

> Students are graded on everything they do every time they turn around. Grades generate anxiety and hard feelings between [everyone]. Common sense suggests that [grades] ought to be reduced to the smallest possible number necessary to find out how students are getting along toward the four or five main objectives of the program, but teachers keep piling them up like squirrels gathering nuts. (p. 17)

2. **Assessment is multidimensional.** In the typical American press for efficiency, we are prone to look for single indices of the effectiveness of any enterprise. We must resist the temptation for simplicity and insist upon multiple measures of any behavior or trait.

The more measures we have, the more we can trust any given conclusion about the performance of a student or a group. In a sense, repeated measures are the means teachers use to establish something akin to a test developer's criterion of reliability. On the other hand, when we use different measures, each offers the prospect of a new perspective and a new insight. Each examines other facets of the learning or looks at learning in a new context or format. Learning varies across both time and situation; assessments must capture the variability in both.

3. **Assessment should be collaborative.** There are many constituencies who have a stake in school assessment--students, teachers, parents, administrators, school board members, the community, and the larger society. Each audience requires somewhat different information, and each should be involved in helping to shape the assessment agenda. At the administrative level, for example, school board members are interested in how well local children read and write in comparison to students across the county, state, school, or classroom. Teachers want to know how well their students are progressing to meet the goals of the school or classroom curriculum; those data will help them plan new instructional steps. Students need to know what is expected of them, how achieving those expectations will help them satisfy personal goals, and their progress in meeting those expectations. To the extent that they can, teachers should take advantage of these perspectives and pool available information to help us all make sound decisions.

The essence of the collaborative criterion lies in collaboration with students. When we work *with* students in developing assessments, we communicate our support of their learning process. In traditional assessment situations we act as passive examiners, detached observers of the "truth" rather than supportive advocates for our students. Collaborative assessments strengthen the bond between teachers and students, and more importantly, allow students to learn about the processes and criteria used in evaluation.

Involving students in assessment moves them from a state of reliance on others to a state of independence in making judgments about their own literacy abilities. In collaborative settings students have opportunities to observe the teacher as evaluator and to examine the criteria and the processes the "expert" uses in evaluation, a modeling strategy similar to that used in apprenticeship training (Collins et al., 1989). And, as students engage in self-evaluation, they are forced to re-examine and refine their own criteria for expertise in reading and writing. By contrast, students who spend their entire school careers subjected to the evaluation criteria of others are not likely to blossom suddenly into self-evaluators.

4. **Assessment must be grounded in knowledge.** Those who accept the responsibility of being "assessors" must be knowledgeable about the content and processes they are assessing. In the case of literacy, they should be familiar with both the basic and instructional processes in reading, writing, and language. At times, we enter into an assessment situation with prespecified criteria and tasks in mind. At other times, we are simply knowledgeable observers of students in their natural learning environment; we look for patterns that will enlighten our understanding and add to our assessment portfolio. And at other times, we interact with students, prompting them, guiding them, questioning them, and listening to them with the skills of a knowledgeable educator. If we want to capitalize on all the assessment possibilities available, there is no substitute for knowledge. The farther we move away from prespecified assessment tasks (e.g., someone else's tests), the greater our need for knowledge of both reading and writing processes and assessment strategies.

5. **Assessment must be authentic.** Authenticity is a requirement that arises from functional learning and ecological validity. Just as instruction should focus on learning to achieve genuine objectives, so too should assessment be anchored in tasks that have genuine purposes (see Edelsky & Draper, in

preparation; Edelsky & Harman, 1988). The application of a literacy task will change depending on these purposes. For example, we might want to assess students' ability to read for details when reading directions, but we would be less concerned about this skill when reading a novel. Mechanics of writing would be of little, if any, concern in journal writing but would have high priority in writing a letter to a public official.

Authenticity of assessment has strong implications for the assessment-learning relationship. On the one hand, assessment tasks should be similar to learning tasks; students should not be surprised by them. On the other hand, assessment tasks should offer students an opportunity to apply, rather than simply regurgitate, knowledge they have acquired. For example, the best index of spelling prowess might be whether a student spells a set of words correctly in a daily assignment, rather than on Friday's test. By allowing assessment practice to accurately reflect learning objectives and behaviors, students will begin to understand that learning is indexed by how they use reading and writing, rather than by a high score on an isolated test. As we move toward authenticity, it is imperative for us to reexamine the purpose and value of any assessment activity and question using any assessment task that asks a student to do anything that has not been a part of a regular learning activity. Authenticity also requires us to worry about the intrinsic value of any assessment activity. And it forces us to adopt a strict, but essential, principle: We should never ask students to do anything on an assessment task that we would not ask them to do as a part of a regular learning activity.

Given these five tenets, we can do no less than support a *portfolio* approach to assessment (Burnham, 1986; Camp, 1985; Carter & Tierney, 1988; Lucas, 1988a, 1988b). We use the term portfolio in both its physical and philosophical senses. As a concept, the "idea" of a portfolio forces us to expand the range of "things" we would consider as data to use in instructional decision-making. It is also an attitude suggesting that assessment is a dynamic rather than a static process reflected most accurately in multiple snapshots taken over time. All who have a stake in contributing to the portfolio (students, teachers, and parents) have a right to "visit" it whenever they wish--to reflect on and analyze its contents. But it is also a physical entity, a *container* in which we can store the artifacts of our literacy--the examples, documents, and observations that serve as evidence of growth and development--so that they are available for visitation and reflection.

A Framework for Conceptualizing Literacy: The Basis for Assessment

Reading and writing can be conceptualized as holistic processes that are enabled by underlying skills. In reading, these could be specific decoding, vocabulary, or comprehension skills. In writing, they could be spelling, punctuation, grammar, usage, rhetorical or organizational skills. But we can also think of reading and writing as processes that serve certain functions in the life of any individual; individuals can read and write to learn information presented, to demonstrate understanding, to gain insight, to communicate with others, to satisfy personal needs, to participate in a culture, or to demonstrate to authorities the trappings of literacy.

Figure 1 illustrates the relationships among the processes of reading and writing, the functional contexts in which these processes exist, and the subskill infrastructure of each as a set of concentric circles. The outer circle represents functional contexts--the purposes literacy serves in our lives. The middle circle represents the processes themselves, intact and undivided. In the innermost circle are all those skills that constitute the infrastructure, or component subskills, of each process.

[Insert Figure 1 about here.]

Infrastructures. Like scientists trying to understand subatomic structure, our assessment focus in the past 20 years has been on that inner circle, those skills constituting the infrastructure of reading and writing. For example, most basal reading programs assess between 15 and 30 specific skills per year in the middle school years (Foertsch & Pearson, 1987). These skills range from knowledge of the meanings of specific words in a reading unit, to root word or affix knowledge, to finding main ideas for

brief passages, to determining the literal meaning of figurative expressions. Students are often asked to take tests in which their writing ability is indexed by their ability to find and correct errors in spelling, grammar, usage, or punctuation. Even students who participate in process writing approaches are sometimes evaluated primarily on strict adherence to the conventions of writing or to some very specific notion of theme organization.

We do not argue that these skills are unimportant or even that they should not be taught or assessed. Rather, we believe there are dangers in perceiving the mastery of such skills as a major instructional goal. Assessing them in isolation leads to highly problematic and misleading conclusions about student performance (Valencia & Pearson, 1988). These "enabling skills" or strategies are important only in that they provide one with the *means* to build a *meaning* for a text. Even strategies such as notetaking, summarizing, self-questioning, predicting, semantic mapping, outlining, and more extended forms of writing are essential only because they provide students with the means to learn from texts (Moore & Murphy, 1987; Tierney, 1982). It is a student's ability to orchestrate and apply these skills in a meaningful context that merits our evaluation. Assessment of reading and writing should focus on the developing expertise needed to apply and adapt strategies and knowledge to many situations (Valencia & Pearson, 1988); that is, skills are best evaluated when they are serving the processes or functions depicted in the two outer layers of the circles in Figure 1 rather than as separate components of the infrastructure.

Holistic processes. The recent curricular refocusing in both reading and writing has moved the assessment focus from subskills to the holistic processes of reading and writing (the middle ring in Figure 1). The assessment emphasis here is on comprehending and composing larger pieces of text in a more global way--on the types of processes we engage in *while* reading and writing. Evidence of this shift first became apparent in 1966, when the National Assessment of Educational Progress implemented holistic scoring of writing samples, attempting to concentrate attention on the attributes of the piece as a "whole" communicative package rather than on the smaller independent components. The comparable starting point for a holistic perspective in reading is probably the popularization of miscue analysis of oral reading errors (Goodman, 1969; Goodman & Burke, 1972); this perspective has been bolstered by the growing use of retellings as a measure of comprehension (e.g., Goodman & Burke, 1972; Mitchell & Irwin, 1988; Morrow, Gambrell, Kapinus, Marshall, & Mitchell, 1986). More recently, the statewide reading assessment efforts in Illinois and Michigan reflect attention to reading comprehension as a holistic understanding of the major ideas and concepts in a text (Valencia & Pearson, 1986; Wixson, Peters, Weber, & Roeber, 1987).

Other holistic assessment efforts can be directed at the processes or strategies involved in reading and writing. For example, we can examine students' ability to monitor their reading and writing as well as their disposition to revise, envision an audience for this writing, or seek help when problems arise (see Baker & Brown, 1984; Garner, 1987; Tierney & Pearson, 1983; Tierney, Soter, O'Flahavan, & McGinley, in press).

Contextualized processes. While holistic assessment perspectives represent a concern for the global processes and an important shift away from isolated skills, they neglect to view reading and writing as processes in communicative contexts (Halliday, 1975; Heath, 1983; Freeman & Sanders, 1987; O'Flahavan & Tierney, in press). We have yet to devise many assessment techniques that emphasize the many different purposes and functions of reading and writing (Applebee, 1984; Freeman & Sanders, 1987; Martin et al., 1975; McGinley & Tierney, 1988). For example, the summaries we write about a novel in school testing situations are likely to be very different from those we create as a part of an informal conversation with friends. Yet we tend to assess summaries per se, without accounting for audience and purpose. Nor have we looked at the interactions among the constructs in three concentric circles in Figure 1. How, for example, we might assess grammar or usage differently when students write to satisfy personal needs as opposed to when they write to request information from someone in a position of authority. This is the essence of what we referred to earlier as "situated" assessment.

Interestingly, some of the more recent critical approaches to literacy have reconceptualized the "enabling skills" depicted in our innermost circle to include the very processes of reading and writing themselves (our middle circle). They view the processes as the means by which one can achieve specific goals both in school and in daily life (Freeman & Sanders, 1987; McGinley & Tierney, 1988; Walters, Daniell, & Trachsel, 1987). Mackie (1981) points out "to be literate is not to have arrived at some predetermined destination, but to utilize reading, writing, and speaking so that our world is progressively enlarged" (p. 1). The traditional notion of "enabling" becomes extended as we move from the "generic" processes of reading and writing to more contextually defined goals of reading and writing.

Interrelationships amongst the circles. If we construe literacy in this way, then reading and writing represent the means or skills with which students can achieve such goals as personal expression, enjoyment, learning, understanding, insight, communication, survival, and cultural participation. Accepting reading and writing as enabling skills suggests that our outermost circle ought to prevail in developing assessment strategies. In the best of situations, we will always assess the processes and their infrastructures as they serve these important literacy functions. When we learn how to account for these functions in our assessment schemes, we will have achieved the kind of authentic, situated assessment inspired by our five attributes of good assessment.

Building an Assessment Portfolio

Most teachers have access to norm-referenced test scores, criterion-referenced test results, and assessments based upon reading and writing activities in the classroom. These are valuable pieces in the assessment puzzle. Norm-referenced tests help us understand students' performance in comparison to other groups of students in the nation; as such, they permit schools and our society to perform important "gate-keeping" and "monitoring" functions. Criterion-referenced measures compare students against some prespecified standard (e.g., 80% correct), usually set by the school, the district, or often a publisher. These forms of assessments offer *one* type of systematic information to add to a portfolio, but they represent a narrowly defined conception of reading and writing.

Too often teacher-constructed assessments, which have the potential to move beyond this narrow conception, seem to be shaped by the standards of commercially produced tests. Studies that have examined textbook and teacher-made tests have indicated that many of them appear frighteningly similar to the ubiquitous norm-referenced test (Armbruster, Anderson, Bruning, & Meyer 1984; Armbruster, Stevens, & Rosenshine, 1977; Foertsch & Pearson, 1987). Even when students are encouraged to write their own answers, they are usually asked to provide only brief, fill-in-the-blank types of responses to lower level questions (Applebee, 1984). Given the potential for variety in classroom assessment, it is remarkable that the "imprint" of standardized tests remains so strong. Teacher-initiated assessments, especially those that deviate from the paper/pencil tradition, have been denigrated as unreliable, subjective, and unfair. Unfortunately, this tight definition of assessment tends to undermine the potential for classroom teachers to act as instructional decision makers (Calfee & Heibert, 1988; Pearson & Valencia, 1987) and leaves no room for creative approaches or professional teacher judgment. But we cannot afford to adopt such a narrow approach to assessment. As Stanley and Hopkins (1972) remind us:

> Let us not fall into the trap of asking whether we should use teacher judgments or test scores. Faced by the complex problems of measurement and evaluation of pupil growth and influences affecting it, we cannot reject any promising resource. Various sorts of information supplement each other. (p. 5)

We maintain that effective assessment must reflect our current understanding of literacy learning. Furthermore, we must simply admit that teacher judgment is an inevitable and important source of evaluative information. We have arrived at a theoretical "crossroads" in our attempt to develop a viable

assessment system. Fortunately, the choice about which road to follow is clear; we must align and integrate assessment and sound instruction in order to prevent the kind of schizophrenic professional life teachers are forced to lead when they switch back and forth from open and flexible instructional practices to closed and rigid assessment systems (Tierney & McGinley, 1987).

However, while the *portfolio* choice is clear, the road to more enlightened assessment practices is neither well paved nor well marked.

A Language for Describing Portfolio Assessment Strategies

As suggested earlier, we are advocating a grass roots portfolio approach to assessment. But this approach is neither lassez-faire nor structureless. In fact, we believe that when teachers leave the security of commercially produced assessments that are conceptualized, monitored, and controlled by "others," what they need are some criteria for classifying, monitoring, and evaluating their own assessment strategies. We have attempted to conceptualize these criteria as a set of five continua on which any approach to assessment can be classified. We have chosen the continuum metaphor in the belief that shifts from one level to another on most of our criteria are more a matter of degree than kind. And, more importantly, these continua will encourage us to explore the range of possibilities and to strive for variety and balance in our approaches to literacy assessment. Our five continua are:

1. focus
2. structure
3. mode
4. locus of control
5. intrusiveness

Focus. The first step in the assessment process is to identify the questions to be addressed. In other words, "What do we hope to learn about this student, group, or class as a result of engaging in this assessment task?" It's helpful to think of the range of questions we might want to ask as spanning the layers of the concentric circles in Figure 1--functional context, holistic process, and infrastructure. Such questions, we maintain, help us to *focus* assessment efforts.

We begin with questions related to the functional context and then move deeper into our circles--the context determines the first line of assessment criteria and also determines how far into the processes and infrastructure one needs to go. While these questions are posed from the perspective of the teacher, they are equally useful to students who wish to become skilled in self-evaluation.

Functional context questions require us to ask how well the intended purpose of the literacy activity is achieved. For example:

How do students use reading and writing for

personal expression?
gaining new insights?
demonstrating knowledge?
learning new information?
getting along in society?
participating in the culture?
demonstrating literacy competence?
thinking critically?

How well are students able to determine and use appropriate forms of reading and writing for a variety of different purposes?

Other questions focus on holistic processes of reading and writing. They require us to assess the integration and application of many subskills and strategies. The focus is on overall meaning and processes rather than on the acquisition of expertise in any specific skills. For example:

How well are students able to use strategies of the composing process such as planning, drafting, revising and editing?

What strategies do students use for sharing writing and getting feedback?

How well are students able to communicate ideas through their writing?

What strategies do students use for getting help with their reading and writing?

How well are students able to guide and monitor their own learning?

How well do students understand the important ideas of a piece of writing?

How well are students able to use strategies of the reading process such as preparing, monitoring, adapting and responding?

Infrastructure questions move us toward the "skills" that enable reading and writing. Although they are not usually a primary concern in the middle school, they still occupy an important place in the curriculum and in the minds of teachers and administrators. They continue to form the foundation for the development of more sophisticated literacy skills and as such, may merit assessment. The key, however, is to assure that these "skills" are assessed in functional contexts, not in isolation. For example:

How well are students able to use the mechanics of writing (e.g., punctuation, capitalization, usage, spelling) to construct meaning?

How well do students organize and sequence ideas in their writing?

How well are students able to use word identification, vocabulary, and comprehension skills to gain meaning from texts?

How well are students able to apply study skills to understand content area material?

We may want to focus on one or several questions during any assessment opportunity. However, this continuum helps to remind us that it is neither necessary nor desirable to hold students accountable for all possible skills and strategies every time we assess. In some cases, we might be concerned with evaluating audience-appropriate writing rather than editing for punctuation. In other cases, we might want to evaluate how well students are able to use context to determine the meanings of several new conceptually important words presented in a science chapter. At other times, we might emphasize the effectiveness of a piece of persuasive writing *and* how well it is organized. Our aim should be to provide a cross-section of assessment foci--a well-balanced assessment portfolio.

Structure. The structure of the assessment indicates how standardized and prespecified the assessment strategy is. Assessments run the gamut from those with specifically defined outcomes and strict guidelines for administration and scoring, to those that are completely unstructured and spontaneous.

structured	semistructured	spontaneous

The most structured assessments specify everything: the desired outcomes/goals, the correct responses, the method of collecting and scoring the data, and any special required time limits or directions. Standardized tests, uniform writing sample assessments, basal reader tests, and "informal" reading inventories are examples of these types of measures. What qualifies these as most structured is the limited amount of input and judgment called for from the classroom teacher and/or the student and the tightly circumscribed types of student responses. Essentially, any trained adult could administer, score, and report results on any of these measures. A slightly less formal measure might be a teacher-constructed multiple-choice test where the teacher designs the content, but leaves other aspects of the assessment to external control.

Somewhere in the middle of the continuum are semistructured, or informal, assessments which require more input and interpretation from the teacher and/or provide greater latitude in students' responses. Asking students to retell a passage just read or to submit an example of a persuasive piece of writing are examples of informal assessment strategies. In the former, the passage is specified but student responses are more open, as are the criteria for evaluating response quality. In the case of the piece of persuasive writing, the assessment is imposed, but the topic choice is open to the student and the scoring criteria can be quite flexible.

At the spontaneous end of the continuum are those unplanned observations, interactions, or examples of student work that provide important insights about how the student is progressing. What distinguishes these assessment activities from the daily, on-the-spot, spontaneous assessments teachers continually make during instruction is nothing more than their inclusion in the portfolio. Spontaneous assessments, whether they are included in a portfolio or not, are important sources of information for instructional decision making; but portfolio documentation provides us with reminders that will help jog our memories and allow us to examine patterns of behavior over time.

Mode. The modes of assessment refer to the processes used to collect information about students' progress. They range from the most *durable* samples of student work to interviews to the less concrete, but no less important, observations.

samples	interviews	observations

We typically think of assessment as pencil and paper tasks or samples of work through which students demonstrate knowledge. These are direct measures of durable products (Cooper, 1981) or, if you will, artifacts of learning. These indices are usually easy to administer or collect, and they provide us with concrete evidence to share with others. They may be initiated by the teacher or by the student, and may be planned or spontaneous. And, they may represent a wide range of foci from literacy functions to subskills. Although many perceive samples of student work as objective, easily evaluated measures, this is not necessarily the case. This misconception probably derives from our naive association of assessment with the term, "test." In reality, the range of possibilities is far greater than we usually imagine.

For example, traditional thinking suggests that this mode of assessment is simply comprised of some type of teacher-constructed written test (multiple-choice, short answer, essay, etc.). A somewhat different type of work sample might be collected for the portfolio by asking students to review several of

the pieces in their writing folders in order to select one they feel is a good example of their writing ability. Another sample might be an audio tape of a student's oral reading or a video tape of a class debate or dramatization. Each of these durable records is an actual sample of student work that can be examined and reexamined over time by teachers, students, and parents.

Although there are attractive advantages to this mode, it is not flawless. Samples of work, by definition, are products; as such, they rarely capture the energy that went into creating them. Additionally, although they possess the lure of objectivity, they are as prone to subjectivity as more "judgmental" forms of assessment. For example, the emphasis on outcome-based learning in the late 1970s led many teachers and school districts to develop their own criterion-referenced assessments. Criterion-referenced assessments usually require a mastery cut-off score, such as 80% correct. What teachers soon found, as they wrote their own tests, was that they had to adjust both text and item difficulty on an ad-hoc basis in order to keep close to their 80% cut score. But because there is no clear conception of task or text difficulty (the knowledge base required in Principle 4), then objectivity is a misnomer. The point of this example is not that subjectivity or teacher judgment are undesirable, rather it is that in paper and pencil tests the subjectivity may simply be moved back one level from public inspection so it is not quite so obvious. Samples of work are not inherently more objective than any other assessment; their value rests in their *durability* over time, not in their "objectivity."

Observations offer several advantages. They permit the evaluator to view learning without intruding on the learning experience. Because they occur more naturally, they can contribute a more valid picture of a student's abilities. At the very least, they afford an opportunity to assess processes and strategies while students are actually learning (see Clay, 1985; Taylor, Harris, & Pearson, 1988).

Observations often arise out of a continuous, less structured approach to assessment. For example, the teacher might observe Henry as he asks his neighbor to help him answer the questions at the end of a social studies chapter, or Susan as she struggles at the writing center for 30 minutes attempting to begin a first draft, or Marcus as he talks to a friend about a book he is reading.

Observations have been criticized for their unreliability, and certainly judgments based on single observations merit such criticism. But this problem can be counteracted in several ways. The first and most important deterrent to unreliability is knowledge. If we know what we are to evaluate, we will evaluate more consistently. This knowledge base can be strengthened and supported by providing teachers with guidelines, checklists, and characteristics to "look for." Second, reliability improves with repeated observations, which have the added advantage of reminding us to look for patterns rather than for single indicators of performance.

Observations have also been criticized as subjective. In response to such criticism, many professionals have advocated tightly planned, structured, systematic observations (Cunningham, 1982). But, as we have argued, subjectivity (in the sense of teacher judgment) is an inescapable characteristic of all assessments.

Using interviews as part of the assessment process enables teachers and students to interact in collaborative settings, share responsibility in shaping the focus of the assessment, and alter the assessment situation through negotiation. Interviews guided by clear purposes but open questions, allow students to share their own views about school, about reading and writing as personal activities, and about how they solve (or fail to solve) their literacy problems. Interviews, be they formal, with a preplanned set of questions, or informal, such as a conversation about a book a student has finished, are often overlooked in our assessment repertoires.

Locus of Control. The responsibility for evaluation usually rests with the teacher, but, as we suggest in Principle 3 (collaboration), students can learn to evaluate their own progress.

teacher collaborative student

The levels on the continuum range from traditional teacher evaluation to collaboration between student and teacher (or between student and peer) to student self-evaluation.

Intrusiveness. The intrusiveness continuum is not independent of the others. Logically, one would expect it to covary with aspects of structure, mode, and locus of control, although the relationship is not linear. An observation checklist that is used as part of the school report card may be quite formal and teacher controlled; however, it would be fairly unintrusive from the student's perspective. But, a two-page book report, required at any time during the semester, might be more informal and offer more shared control, and yet be quite intrusive.

intrusive unintrusive

Some Illustrative Scenarios

Because literacy assessment is so complex, it is impossible for us to recommend an exhaustive set of assessment strategies for classroom use. The very reason that we have described our principles and criteria (our continua) so carefully is that we believe teachers must tailor assessment practices to their classroom situations. Guidelines, not tests, will serve them better in this endeavor. To illustrate how this can be done, we provide some illustrative assessment scenarios. For each scenario, we will depict the specific assessment questions and procedures used by a middle school teacher to evaluate some aspect(s) of reading and writing performance. Then we will classify each according to our five continua and elaborate ways in which the strategy might be adapted to answer different questions or fulfill different assessment needs. Each scenario has these subsections: *The question* (the assessment question the teacher was trying to answer); *The situation* (a description of the context and the assessment activities used); *Rating the assessment activity* (here we indicate where, on each of the five continua, we would place the activities); *What can we learn* (an explanation of how the activity might be adapted to other situations).

Scenario 1

The question. How well are students able to guide and monitor their own learning?

The situation. A seventh-grade English teacher wants to help his students improve their ability to monitor their own reading to insure that it makes sense to them. He has his students use paper bookmarks to record problem spots (e.g., difficult concepts, unknown words, questions they would like to discuss) they encounter while reading the assigned chapter in a novel. After they discuss the chapter in class, he holds an individual conference with each of several students (different students are scheduled for conferences each week). First he has the student retell the chapter to check on comprehension (see Appendix A). If the student has read with good understanding and has not identified any problem areas, the teacher decides not to use this opportunity to assess self-monitoring. If the student has exhibited some comprehension difficulty and/or has identified possible problems while reading, then they discuss the trouble spots. They review the notes on the bookmark, review the chapter for other potential problem spots not noted, and then discuss strategies for coping with the difficulties. The teacher jots down some anecdotal notes and sometimes even attaches the student's bookmark to his notes for later placement in the student's portfolio.

Rating the assessment activity

focus	infrastructure	X holistic	function
structure	structured	X semistructured	spontaneous
mode	X samples	X interview	observation
locus of control	student	X collaborative	teacher
intrusiveness	X intrusive		unintrusive

What can we learn? This task amounts to asking students to "edit" their own reading and writing. In reading, because most middle school students don't have the luxury of writing in their texts while they read, we might use this bookmark technique. By observing their notes, interviewing, or even collecting these bookmarks, teachers may gain insight into each student's ability to process a piece of text. Collaboratively student and teacher can identify possible problems and design instructional strategies.

In writing, students might be asked to proofread (edit) their drafts with one or two specific purposes in mind. These purposes may be as detailed as proofreading for spelling or punctuation errors or as broad as revising the presentation of a line of reasoning or an argument. The student may then confer with the teacher (or submit papers), using both the first and the second drafts to evaluate progress collaboratively.

This particular type of activity can be framed in a variety of ways. The focus may be on specific vocabulary skills, author's craft, prewriting strategies, or organization; the mode could be observation interview, or sample; the activity may be teacher initiated or an unintrusive student initiated one. Thus, it has the potential to fall virtually anywhere along each of the five continua. Furthermore, this assessment activity can be repeated using different material and focusing on many different aspects of reading or writing.

There are many other strategies for gathering information about students' ability to monitor their own learning. Another way to assess self-monitoring is simply to ask them to annotate each work sample as they file it in their portfolio. They might rate its quality on a general 1-10 scale, they might rate it on specific criteria, or they might write a short statement about the process, difficulties encountered, or their degree of satisfaction with the product.

Scenario 2

The question. How do students respond to a piece of literature?

The situation. An eighth-grade English class has just finished reading *The Diary of Anne Frank*, and the teacher has asked them to write an essay of no more than three pages detailing their reaction and response to the book in a way that might provide a helpful overview to another group of eighth-grade students who are about to begin reading the book. As she finishes her directions for the assignment, the teacher reminds the students to focus on the content of the responses and not to worry about mechanics and grammar on this early draft; there will be plenty of time for that, she tells them, when they complete the version they will *actually* send to students in her other eighth-grade English class.

Rating the assessment activity

focus	infrastructure	holisitic	X function
structure	structured	X semistructured	spontaneous
mode	X samples	interview	observation
locus of control	student	collaborative	X teacher
intrusiveness	X intrusive		unintrusive

What can we learn? For the first draft, the teacher is clearly interested in evaluating students' response to a piece of literature rather than their ability to control the mechanical features of the writing process. She could use a holistic scale (see Appendix B) for evaluating the writing activity. A holistic scale places the focus on the quality of the communication to the potential audience. Alternatively, she could use an analytic scale (see Appendix C for an example) in which she had separate subscales for traits such as fidelity to the themes of the book, sense of audience, organization, and style. When the students complete the version they will send, she may want to incorporate subscales dealing with the conventions of written language and/or the "impact" of the message. But notice that the motivation for adhering to conventions is inherently greater when the students have a real audience in mind rather than audiences consisting of the teacher or some vague, hypothetical group ("Pretend you are writing to . . .").

Although this activity is primarily a writing activity, the teacher can learn something about her students' reading performance. For example, what the students regard as important enough to relate to another student is a good index of what they attended to when reading the text.

As in most of our other illustrative scenarios, with slight modifications, this scenario could take on many different values on most of the five continua. For example, if the teacher tells students that correct mechanics are absolutely essential, she can shift the focus of the activity from the context to the infrastructure immediately. Or, if she specifies the exact form of the response (e.g., a five-paragraph theme), then the task becomes very structured and even more intrusive.

Scenario 3

The question. How well are the students able to compose an informational text?

The situation. A sixth-grade language arts teacher requests that students keep a portfolio of all the writing they do throughout the year. As a result, each student's portfolio contains several samples of narrative, personal, and informational writing. At various points throughout the year, she asks her students to select a piece which they feel represents their best effort at composing in a particular genre. Prior to submitting their paper, students are given the opportunity to make any final revisions they see as necessary.

Before they submit it, she informs them that the piece must be accompanied by a letter to her explaining and justifying its selection. She also tells the students that she will evaluate the letters on the basis of the three aspects of writing that they have been studying in the last month--quality of the content, organization, and voice. In addition, she adds, she will evaluate each letter on the basis of how

persuasive it is in convincing her that the piece really does include these critical qualities of good informational writing.

Rating the assessment activity

focus	infrastructure	holisitic (X)	function
structure	structured	semistructured (X)	spontaneous
mode	samples (X)	interview	observation
locus of control	student (X)	collaborative	teacher
intrusiveness	intrusive	unintrusive (X)	

What can we learn? In this scenario, the teacher is interested in learning a number of things about her students' ability to compose informational texts. The pieces themselves provide a source of information about their knowledge of the specific content they are reporting, as well as their ability to organize and synthesize information in their own language. Most importantly, by asking students to write a letter justifying the quality of their selection, she has asked that they begin to take steps toward becoming evaluators of their own writing. In doing so, she has created a situation with equal potential for both instruction and assessment. In writing about their writing, students are given an opportunity to evaluate and rethink the standards they associate with good informational writing. By examining similar self-evaluation over time, the teacher can determine those features of writing that are focusing on students and how their focus changes as a result of instruction over the course of the year: Are they concerned with features related to the infrastructure of the piece, the piece as a whole, or with the functions or purposes that it serves?

This activity may be framed in a number of other ways. For example, in asking students to write about their writing, the teacher may choose to tighten the structure by specifying the criteria for evaluation or to open the structure by asking students to select their criteria for evaluation.

Scenario 4

The question. How well are students able to use their prior knowledge and reading skill to learn new information from a textbook?

The situation. In his sixth-grade social studies class, a teacher has decided to use a pre-reading followed by a post-reading group semantic mapping activity as a way of preparing for and then following up the reading of the chapter on the development of Greek civilization. For the prereading mapping activity, he begins with a blank chart and writes the words, *Ancient Greece*, in the middle of the chart, asking students to share ideas that come to mind when they think of Ancient Greece. Once 25 to 30 ideas have been added to the chart (within categories), he asks the students to read the chapter and look for new ideas to add to the chart or think of ways they may wish to revise it. Once the reading is finished, he meets with the whole class to revise the chart--adding another 30 ideas to it, putting in labels for some of the categories that have evolved, and switching a few ideas from one category to another. As a culminating activity, he asks that each student develop a written summary of the information included in the class map.

Rating the assessment activity

focus	infrastructure	X (summary) holisitic	X (map) function
structure	structured	X semistructured	spontaneous
mode	X summary samples	interview	X (map) observation
locus of control	student	X collaborative	teacher
intrusiveness	intrusive	X	unintrusive

What can we learn? This activity is interesting because other than the summary at the end of the whole activity, there is no provision for any artifacts representing the performance of individuals; instead, the map represents a group effort. During the mapping activity itself, the teacher has many opportunities to observe the contributions of individual students; he can draw conclusions about their knowledge before reading the chapter as well as what they learned while reading it. He can also make anecdotal notes to add to the portfolio. The summary activity provides the teacher with a chance to learn a great deal about the students' ability to organize and present a set of ideas.

With some modification, mapping can provide more insight into individual performance. For example, some teachers ask students to complete the pre-reading and post-reading maps on their own. In such cases, the maps become individual rather than group artifacts. The cost of individual mapping is that students do not get a chance to learn from one another during the group mapping sessions. As a more intrusive measure (but perhaps more informative with respect to assessing growth for particular concepts presented in the text) the maps can also be much more structured than in this example. The teacher can provide all the categories for the association task, in which case it may look more like a study guide or an incomplete outline.

Scenario 5

The question. How well are students able to use reading and writing to create a research report?

The situation. A ninth-grade language arts teacher has just completed a month-long instructional unit on report writing. She initially met with the science teacher to discuss what topics the students would be studying in November. Together they outlined possibilities for a related language arts writing unit. Students were required to (a) select a topic of personal interest related to the solar system, planets, or space exploration; (b) state a hypothesis; (c) gather information from at least three sources; and then (d) develop a complete report to be shared with their peers as well as their language arts and science teachers.

Over the course of several weeks, the students were introduced to several different strategies for selecting and narrowing topics of study. In addition, the teacher has introduced the students to many possible sources of information in the library (e.g., encyclopedia, books, magazines, films, etc.) and in the community (e.g., interviews with experts, visits to museums, etc.).

Because the teacher had been using a process approach to teaching writing, she helped students use the strategies they had acquired to compose their reports. Additionally, she set aside time over several weeks to work with students on two new skills--synthesizing information from several sources and using

footnotes and bibliographic information. Students shared their final reports during two combined language arts and science class periods. Students with common interests formed groups in which they shared and discussed their reports and provided feedback. The reports were submitted to both teachers for grading.

As students moved through their drafts and revisions, the teacher observed the strategies they used for getting started and for getting help with revision. Using a checklist several times during the month as a guideline (see Appendix D for an example), she kept track of each student's progress and used the data to form short-term instructional groups on several of the goals and strategies explicit in the unit.

Rating the assessment activity (the checklist)

focus	infrastructure	X holistic	function
structure	structured	X semistructured	spontaneous
mode	samples	interview	X observation
locus of control	student	collaborative	X teacher
intrusiveness	intrusive		X unintrusive

During the month the teacher also suggested that students write in their learning logs at least twice a week (see Kirby & Liner, 1981, for examples). They were encouraged to write about their learning experiences as they gathered information and composed their reports: their problems, need for additional guidance, new insights, and feelings about their efforts. The logs were shared with the teacher at the students' discretion.

Rating the assessment activity (learning logs)

focus	infrastructure	student determined holisitic	function
structure	structured	semistructured	X spontaneous
mode	samples	X interview	observation
locus of control	X --> --> --> student	collaborative	teacher
intrusiveness	intrusive		X unintrusive

The final report was evaluated using information from four sources: the language arts teacher, the science teacher, the student, and the special interest group with whom the report was shared. Each evaluator considered the report from a different perspective. The language arts teacher used three main criteria: a holistic scoring of the overall quality of the paper, the synthesis of several sources of

information into a coherent original piece, and appropriate form for footnotes and bibliographic information (a primary instructional goal for this unit). Students were provided with a self-evaluation questionnaire (see Appendix E) to guide their reflections on the process of writing the report and on the quality of the final product. The science teacher evaluated both the quality of the information presented and students' ability to use a scientific report genre--to state an hypothesis, present information, and draw appropriate conclusions. The interest group evaluated each report on how well the student presented the information to the group and the author's ability to react to questions about the report. All these assessments were used to arrive at a final grade for the project.

Rating the assessment activity (the report)

focus	___X_____	____X_____	____X____
	infrastructure	holisitic	function
structure	__X_____		
	structured	semistructured	spontaneous
mode	_____X_____		____X____
	samples	interview	observation
locus of control	__X_->_->_->	_____X_____	____X__
	student	collaborative	teacher
intrusiveness	_____X_____		
	intrusive		unintrusive

What can we learn? This rich instructional context provides the teacher with unlimited opportunities to assess students during the entire month-long process of reading and writing to create their reports. The instructional unit also results in a concrete artifact--the report--and the added experience of presenting the work to peers and to experts. Notice that for many of our continua, we have marked several points rather than a single point. This is because in situations such as these, assessment opportunities tend to be flexible and dynamic; that is, a single strategy can relate to more than one structure, mode, or level of focus, and it can change over time.

Conclusion

In this report, we have tried to build a strong case for a portfolio assessment for reading and writing during the middle school years. A portfolio is both a physical collection of the artifacts of students' reading and writing and a disposition toward assessment. This disposition is characterized by the belief that assessment must be continuous, multidimensional, collaborative, knowledge-based, and authentic. Our recommendation for portfolios is based upon the conviction that records gathered while students are engaged in functional and contextualized literacy tasks will ultimately prove more useful to both teachers and students than will any set of numbers derived from tests that have little relevance to the purposes and needs of either students or teachers.

The scenarios and the assessment tools we have provided are illustrative rather than exhaustive, descriptive rather than prescriptive. The principles and continua put forward should be regarded as a conceptual framework to use in developing and evaluating assessment strategies. The array of assessment opportunities available to us as teachers is limited only by our creativity, our knowledge of the reading process, and our grasp of sound assessment criteria. It is our responsibility to avail ourselves of these opportunities.

References

Anderson, R. C., Hiebert, E. H., Scott, J. A., & Wilkinson, I. A. (1985). *Becoming a nation of readers*. Urbana-Champaign: University of Illinois, Center for the Study of Reading.

Applebee, A. N. (1981). *Writing in the secondary school: English and the content areas*. Urbana, IL: National Council of Teachers of English.

Applebee, A. N. (1984). *Contexts for learning to write*. Norwood, NJ: Ablex.

Armbruster, B. B., Anderson, T. H., Bruning, R. H., & Meyer, L. A. (1984). *What did you mean by that question: A taxonomy of American history questions* (Tech. Rep. No. 308). Urbana-Champaign: University of Illinois, Center for the Study of Reading.

Armbruster, B. B., Stevens, R. J., & Rosenshine, B. (1977). *Analyzing content coverage and emphasis: A study of three curricula and two tests* (Tech. Rep. No. 26). Urbana-Champaign: University of Illinois, Center for the Study of Reading.

Baker, L., & Brown, A. L. (1984). Metacognitive skills and reading. In P. D. Pearson, R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research*. New York: Longman.

Burnham, C. C. (1986). Portfolio evaluation: Room to breathe and grow. In Charles Bridges (Ed.), Training the teacher of college composition. Urbana, IL: National Council of Teachers of English.

Calfee, R., & Hiebert, E. (1988). *The teacher's role in using assessment to improve learning*. Unpublished manuscript.

Camp, R. (1985). The writing folder in post-secondary assessment. In P. J. A. Evans (Ed.), *Directions and misdirections in English evaluation*. Ottawa, Canada: The Canadian Council of Teachers of English.

Carter, M. A., & Tierney, R. J. (1988). *Reading and writing growth using portfolios in assessment*. Paper presented at the annual meeting of the National Reading Conference, Tucson, AZ.

Clay, M. M. (1985). *The early detection of reading difficulties* (3rd ed.). Auckland, New Zealand: Heinemann.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics. In L.B. Resnick (Ed.), *Knowing, learning & instruction: Essays in honor of Robert Glaser*. Hillsdale, NJ: Erlbaum.

Cooper, J. O. (1981). *Measuring behavior*. Columbus, OH: Charles E. Merrill.

Cunningham, P. M. (1982). Diagnosis by observation. In J. J. Pikulski & T. Shanahan (Eds.), *Approaches to the informal evaluation of reading*. Newark, DE: International Reading Association.

Diederich, P. (1974). *Measuring growth in English*. Urbana, IL: National Council of Teachers of English.

Edelsky, C., & Draper, K. (in preparation). *Authenticity: A purposeful notion*.

Edelsky, C., & Harman, S. (1988). One more critique of reading tests--with two differences. *English Education, 20*, 157-171.

Faigley, L., Cherry, R. D., Jolliffe, D. A., & Skinner, A. M. (1985). *Assessing writers' knowledge and processes of composing*. Norwood, NJ: Ablex.

Foertsch, M., & Pearson, P. D. (1987). *Reading assessment in basal reading series and standardized tests*. Paper presented at the annual meeting of the National Reading Conference, St. Petersburg, FL.

Freeman, E. B., & Sanders, T. (1987). The social meaning of literacy: Writing instruction and the community. *Language Arts, 64*, 641-645.

Gage, J. (1986). Why write? In D. Bartholomea & A. Petrosky (Eds.), *The teaching of writing* (pp. 8-29). Chicago: National Society for the Study of Education.

Garner, R. (1987). *Metacognition and reading comprehension*. Norwood, NJ: Ablex.

Goodman, K. S. (1969). Analysis of oral reading miscues: Applied psycholinguistics. *Reading Research Quarterly, 5*, 9-30.

Goodman, Y. M., & Burke, C. L. (1972). *Reading miscue inventory manual: Procedures for diagnosis and evaluation*. New York: Macmillan.

Halliday, M. (1975). *Explorations in the functions of language*. London: Edward Arnold.

Heath, S. B. (1983). *Ways with words*. Cambridge, MA: Cambridge University Press.

Kirby, D., & Liner, T. (1981). *Inside out: Developmental strategies for teaching writing*. Boynton/Cook.

Langer, J. A., & Applebee, A. N. (1987). *Writing and learning in the secondary school*. (National Institute of Education Grant No. NIE- G-82-0027.) School of Education, Stanford University.

Lucas, C. K. (1988a). Toward ecological evaluation. *The Quarterly of the National Writing Project and the Center for the Study of Writing, 10*(1), 1-17.

Lucas, C. K. (1988b). Toward ecological evaluation, Part Two. *The Quarterly of the National Writing Project and the Center for the Study of Writing, 10*(2), 4-10.

Mackie, R. (1981) (Ed.). *Literacy and revolution: The pedagogy of Paulo Freire*. New York: Continuum.

Martin, N., Meday, P., Smith, H., and D'Arcy, P. (1975). Why Write? In N. Martin (Ed.), *Writing across the curriculum*. Montclair, NJ: Boynton/Cook.

McCaig, R. A. (1981). A district-wide plan for the evaluation of student writing. In S. Haley-James (Ed.), *Perspectives on writing in Grades 1-8*. Urbana, IL: National Council of Teachers of English.

McGinley, W., & Tierney, R. J. (1988). *Reading and writing as ways of knowing and learning* (Tech. Rep. No. 423) Urbana-Champaign: University of Illinois, Center for the Study of Reading.

Mitchell, J. N., & Irwin, P. A. (1988). *The reader retelling profile: Using retellings to make instructional decisions*. Unpublished manuscript.

Moore, D. W., & Murphy, A. G. (1987). Reading programs. In D. E. Alvermann, D. W. Moore, & M. W. Conley (Eds.), *Research within reach secondary school reading: A research guided response to concerns of reading educators*. Newark, DE: International Reading Association.

Morrow, L., Gambrell, L., Kapinus, B., Marshall, N., & Mitchell, J. (1986). Retelling: A strategy for reading instruction and assessment. In J. Niles (Ed.), Solving problems in literacy: Learners, teachers, and researchers. *Thirty-fifth yearbook of the National Reading Conference*. Rochester, NY: National Reading Conference.

O'Flahavan, J. F., & Tierney, R. J. (in press). Reading, writing, and critical thinking. In B. F. Jones & L. Idol (Eds.), *Dimensions of thinking and cognitive instruction* (Vol. 2). Hillsdale, NJ: Erlbaum.

Pearson, P. D., & Valencia, S. W. (1987). Assessment, accountability, and professional prerogative. *Research in literacy: Merging perspectives: Thirty-sixth yearbook of the National Reading Conference*. Rochester, NY: National Reading Conference.

Purves, A. (1982). *International evaluation of assessment of written composition*. Unpublished manuscript, University of Illinois at Urbana-Champaign.

Shulman, L. S. (1975). Teaching as clinical information processing. In N. L. Gage (Ed.), *National conference on studies in teaching*. Washington, DC: National Institute of Education.

Stanley, J. C., & Hopkins, K. D. (1972). *Educational and psychological measurement and evaluation*. Englewood Cliffs, NJ: Prentice-Hall.

Taylor, B., Harris, L. A., & Pearson, P. D. (1988). *Reading difficulties: Instruction and assessment*. New York: Random House.

Temple, C., & Gillet, J. W. (1984). *Language arts: Learning processes and teaching processes*. Boston: Little, Brown & Company.

Tierney, R. J. (1982). Learning from text. In A. Berger & H. A. Robinson (Eds.), *Secondary school reading: What research reveals for classroom practice*. Urbana, IL: ERIC Clearinghouse on Reading and Communication Skills and the National Council on Research in English.

Tierney, R. J., & McGinley, W. (1987). *Serious flaws in written literacy assessment*. Paper presented at the annual meeting of the American Educational Research Association.

Tierney, R. J., & Pearson, P. D. (1983). Toward a composing model of reading. *Language arts, 60*, 568-580.

Tierney, R. J., Soter, A., O'Flahavan, J. F., & McGinley, W. (in press). The effects of reading and writing upon thinking critically. *Reading Research Quarterly*.

Valencia S. W., & Pearson, P. D. (1986). *Reading assessment initiatives in the State of Illinois--1985-1986*. Springfield, IL: Illinois State Board of Education.

Valencia, S. W., & Pearson, P. D. (1987). Reading assessment: Time for a change. *Reading Teacher, 40*, 726-732.

Valencia, S. W., & Pearson, P. D. (1988). Principles for classroom comprehension assessment. *Remedial and Special Education, 9*, 26-35.

Walters, K., Daniell, B., & Trachsel, M. (1987). Formal and functional approaches to literacy. *Language Arts, 64*, 855-868.

Winson, K. K., Peters, C. W., Weber, E. M., & Roeber, I. (1987). New directions in statewide reading assessment. *Reading Teacher, 40*, 749-755.

APPENDIX A

A Schema for Scoring Retellings
The Retelling Profile

Directions. Indicate with a checkmark the extent to which the reader's retelling includes or provides evidence of the following information.

	none	low degree	moderate degree	high degree
1. Retelling includes information directly stated in text.				
2. Retelling includes information inferred directly or indirectly from text.				
3. Retelling includes what is important to remember from the text.				
4. Retelling provides relevant content and concepts.				
5. Retelling indicates reader's attempt to connect background knowledge to text information.				
6. Retelling indicates reader's attempt to make summary statements or generalizations based on text that can be applied to the real world.				
7. Retelling indicates highly individualistic and creative impressions of or reactions to the text.				
8. Retelling indicates the reader's affective involvement with the text.				
9. Retelling demonstrates appropriate use of language (vocabulary, sentence structure, language conventions).				
10. Retelling indicates reader's ability to organize or compose the retelling.				

Appendix A (Continued)

	none	low degree	moderate degree	high degree
11. Retelling demonstrates the reader's sense of audience or purpose				
12. Retelling indicates the reader's control of the mechanics of speaking or writing.				

Interpretation. Items 1-4 indicate the reader's comprehension of textual information; items 5-8 indicate metacognitive awareness, strategy use, and involvement with text; items 9-12 indicate facility with language and language development.

Note. From J. N. Mitchell and P. A. Irwin (1988).

APPENDIX B

A Wholistic Scoring Scale

Level 1 . . . Not competent

Either the content is inadequate for the topic selected, or deficiencies in the conventions of written expression are so gross that they interfere with communication.

Level 2 . . . Not competent

The student can express a message which can be readily understood, which contains adequate content for the selected topic, and which demonstrates at least marginal command of sentence sense.

The writing, however, is grossly deficient in one or more of these skills, judged by standards appropriate for high school:

Spelling
Usage
Punctuation and capitalization

Level 3 . . . Marginally competent

The student can compose a completed series of ideas about a topic with a minimum of gross deficiencies in spelling, usage, or punctuation, judged by standards appropriate for high school.

The writing, however, does not contain at least one competent paragraph or is not competent in one or more of these skills, judged by standards appropriate for high school:

Sentence sense
Spelling
Usage
Punctuation and capitalization

Level 4 . . . Competent

The student can compose a completed series of ideas about a topic with basic skills at a level appropriate for high school and with at least one competent paragraph.

The writing, however, does not demonstrate all of the characteristics of highly competent writing:

Good overall organization
Competent paragraphing
Regular use of transitions
Interpretive meaning (as opposed to literal writing)
Good sentence structure
Good vocabulary
Appropriate use of subordination

Appendix B (Continued)

Level 5 . . . Highly competent

The student can compose a completed series of ideas about a topic with basic skills at a level appropriate for high school and with these characteristics of highly competent writing:

- Good overall organization
- Competent paragraphing
- Regular use of transitions
- Interpretive meaning (as opposed to literal writing)
- Good sentence structure
- Good vocabulary
- Appropriate use of subordination

The writing does not, however, demonstrate thesis development and does not contain critical or creative thinking.

Level 6 . . . Superior

The student can compose a completed series of ideas about a topic with excellent basic skills, with the characteristics of highly competent writing, with adequate thesis development, and with at least one passage demonstrating critical or creative thinking.

The passage of superior writing, however, tends to be an isolated example.

Level 7 . . . Superior

The student can compose a completed series of ideas about a topic with excellent basic skills, with critical or creative thinking, and with a sustained vitality and richness of expression.

Note. From McCaig (1981).

APPENDIX C

An Analytic Scoring Scale

A. Quality and Scope of Intent

"5" clear perception of topic, problems, and issues which is outlined in the introduction/seriousness may not be explicit, but is implied/use of detail is specific to issues raised

"3" reader can discern point of view of writer, but not as clearly as 5/no development issues before launching into pros and cons/detail is minimal, glossy

"1" superficial addressing of the issue/"I think" stance/broad, general statements/little or no detail

B. Organization and Presentation of Content

"5" focus of topic is clear/digressions are relevant and usually signalled/all or most of essay is clearly elaborative, supports arguments fully through logical development

"3" lacks clear focus, but remains on topic/some evidence of faulty reasoning, but adequate development of reasoning/sense of beginning, middle, end

"1" has not given reader sense of which aspect is most important, or is incomprehensible/scant or no paragraph development/scant or no elaboration of main points through use of examples, etc.

C. Style and Tone

"5" language generally consistent with rhetorical task and consistent within itself/no gaps in information/generally, content should link smoothly through logical development/variety of sentence types

"3" more appeals to reader's emotions than intellect/no slang, no jargon/little or no use of subordination/reader can follow flow of thought, but less so than 5

"1" language detracts from persuasive task/may be flippant/lack of connectivity or less so than 3/some gaps in content/little or no flow between ideas

Appendix C (Continued)

D. Mechanics

"5" 1 or no errors in usage/generally, very few or no spelling errors or punctuation errors

"3" errors in usage, but text remains comprehensible/errors in spelling and punctuation, but text comprehensibility is not affected

"1" text comprehensibility greatly affected by errors/frequent, flagrant errors, where text is difficult to process

E. Affective Response of the Raters

"5" essay was interesting/a sense of connection with the writer

"3" moderately interesting/a sense of moderate connection with the writer

"1" not really interesting/no sense of connection with the writer (avoid the writer at all costs)

F. Overall Impression Score

"7" off task

"5" exceptional essay or very near exceptional: persuasive argument/use of accepted, mechanical conventions, logically developed, connected prose/rich detail, elaboration/generally, good flow of thought

"3" an acceptable piece: few mechanical errors/more gaps in information than in 5/development of ideas is there, but must be flushed out by reader/an occasional appeal to reader's emotions rather than intellect

"1" a poor effort: many errors in mechanics, consistency of tone, logical development, flow of thought, connectedness/comprehension of the piece adversely affected

Note. Adapted from A. Purves (1982).

APPENDIX D

A Writing Analysis Checklist

The Writing Process

1. How often does the writer get ideas for writing
   - ___ from the imagination?
   - ___ from discussion with others?
   - ___ by imitating a book, story, poem, TV show, and so on?
   - ___ from the teacher's assignments?
   - ___ from some other source? which?

2. When the writer means to rehearse what will be written, and to narrow down the topic, does the writer
   - ___ talk to classmates?
   - ___ talk to the teacher?
   - ___ draw a picture first?
   - ___ think about it overnight?
   - ___ start writing right away?

3. In drafting a paper, does the writer
   - ___ write one draft only?
   - ___ invent spellings, use a dictionary, or limit vocabulary to the words he or she can spell?
   - ___ scratch out words and lines, and cut and paste?
   - ___ seek comments from others about the way the drafting is going?

4. Does the writer revise a paper before it is considered finished? Do the drafts
   - ___ all look like different papers, new beginnings?
   - ___ look like mechanical refinements of earlier drafts?
   - ___ interact with and build on the ideas of early drafts?
   - ___ stop after one draft?

The Functions of Writing

5. What forms of writing has the writer produced?
   - ___ stories?
   - ___ poems?
   - ___ expressive writing (personal experiences and opinions)?
   - ___ persuasive writing?
   - ___ descriptive writing?
   - ___ expository writing (that which explains or gives directions)?

6. What kinds of topics has the writer written about?
   - ___ topics about which the writer was an expert?
   - ___ topics about which the writer had to learn more before writing?
   - ___ topics about things that were present?
   - ___ topics about things that were past or absent?
   - ___ topics about abstract ideas?

Appendix D (Continued)

7. What audiences has the child written for?
   ___ the teacher?
   ___ classmates?
   ___ other people known to the child? Whom?
   ___ other people unknown to the child? Whom?

8. In trying to stick to the topic, did the writer
   ___ limit the focus of the topic before starting to write?
   ___ stick to one thing or ramble?
   ___ focus more on the object of the writing or on the writer?

9. In trying to stick with the purpose of writing, does the writer
   ___ keep expressing personal feelings, although the topic and purpose suggest doing otherwise?
   ___ declare one purpose but pursue another (such as "The story about . . . ," which is expository, not narrative)?
   ___ shift from one purpose to another?

10. In trying to meet the audience's need for information
    ___ does the writer appear to assume the audience knows and is interested in the author?
    ___ is he or she careful to tell the audience things they will need to know in order to understand what is talked about?
    ___ does the writer address the same audience throughout?

Qualities of Writing Style

11. Does the writer use exact, well-chosen words?

12. Does the writer "paint pictures with words"--make the reader see what the writer saw?
    ___ is the focus on immediate, "here-and-now" images?

13. In regard to the organization of the papers:
    ___ does the writer keep the focus on one aspect of the topic at a time?
    ___ do the papers have identifiable openings?
    ___ are the details arranged in a reasonable order and do they relate reasonably to one another?
    ___ is there an identifiable ending to the papers?

Fluency of Writing

14. How long are the papers (in words or lines per paper)?

15. What is the average number of words per sentence?

16. What is the average number of words per T-unit?

Appendix D (Continued)

Mechanics of Writing

17. In handwriting, does the writer
    - ___ have problems forming letters? Which ones?
    - ___ have problems spacing between letters? Keeping vertical lines parallel? Keeping the writing even on the baseline?
    - ___ write with uniform pressure? In smooth or in jerky lines?

18. In regard to spelling
    - ___ does the writer misspell words in the first draft?
    - ___ does the writer correct the spellings of many words between the first and later drafts?
    - ___ what does the writer do when uncertain of how to spell a word?

19. Does the writer have trouble with standard English usage?
    - ___ does the writer write in complete sentences? If not, what are the units of writing like?
    - ___ does the writer have problems with punctuation and capitalization? With which elements? In what circumstances?
    - ___ are errors made in standard English grammar? If so, describe the errors.

Enjoying Writing

20. Does the writer take pleasure in writing?
    - ___ how do you know?

Note. From Temple and Gillet (1984).

APPENDIX E

A Self-Evaluation Questionnaire

1. List the most successful things you did in writing this paper.

   List the things that a reader will think are successful.

2. List the things you were unable to do in this paper that would have made it more successful.

3. In the process of writing this paper, what aspects were easier than when you have written previous papers?

4. In the process of writing this paper, what aspects were more difficult than when you have written previous papers?

Note. From Faigley, Cherry, Jolliffe, & Skinner (1985).

Figure 1. Dimensions of Reading and Writing